



**SADA**

**South African Data Archive**

**GENERAL HOUSEHOLD SURVEY, 2007  
(South Africa)**

**Statistics South Africa (Stats SA)**

**CODEBOOK**

**SADA 0156**

## **BIBLIOGRAPHIC CITATION**

Any publication or other presentation based in whole or part on the data and documentation supplied by SADA must prominently use the following citation:

Statistics South Africa. *General Household Survey, 2007* [Computer file]. S0156. Pretoria: Statistics South Africa [producer], 2007. Pretoria: South African Data Archive, National Research Foundation [distributor], 2008.

## **DEPOSIT OF PUBLICATIONS**

At least one copy of any published work or report based in whole or part on the dataset will be deposited with the South African Data Archive. Please indicate the title and number of the study utilised.

## **LIMITATION OF LIABILITY**

The Archive and the depositor of the dataset supplied to the user bear no legal responsibility for their accuracy or comprehensiveness.

Postal Address:	SOUTH AFRICAN DATA ARCHIVE NATIONAL RESEARCH FOUNDATION P.O. BOX 2600 Pretoria 0001
Telephone:	+27 12 481 4158/ 4016/ 4120
Telefax:	+27 12 349 1179
Electronic mail:	<a href="mailto:sada@nrf.ac.za">sada@nrf.ac.za</a>
World Wide Web:	<a href="http://www.nrf.ac.za/sada">http://www.nrf.ac.za/sada</a>

## STUDY DESCRIPTION

### SADA 0156

**TITLE:** General Household Survey, 2007 (GHS 2007).

**PRINCIPAL INVESTIGATOR:** Statistics South Africa (Stats SA)

**DEPOSITOR:** Statistics South Africa

**ORGANISATION HOUSING THE DATA:**

Statistics South Africa  
Private Bag x44  
Pretoria, 0001  
South Africa  
Tel: (012) 310 8600  
Fax: (012) 310 8500  
Email: [info@statssa.gov.za](mailto:info@statssa.gov.za)  
Web: <http://www.statssa.gov.za>

**KEYWORDS:**

Employment; Unemployment; Household surveys; Housing; Education and Health.

**SERIES DATA:**

Series name:	Principal Investigator:	Currently available:
General Household Survey	Statistics South Africa	2002 - 2011

**PURPOSE:**

The GHS is an annual household survey, specifically designed to measure various aspects of the living circumstance of South Africa households.

**GEOGRAPHIC LOCATION:** South Africa

**IMPORTANT VARIABLES:** Employment, unemployment, education, health, access to services and facilities.

**DEMOGRAPHIC VARIABLES:** Age, gender, level of education, marital status, use of health services, economic activity and employment.

**UNIVERSE:** Households in the nine provinces of South Africa

**METHOD OF DATA COLLECTION:** Survey Questionnaire

**SAMPLING:**

Enumeration Areas (EAs) that had a household count of less than twenty-five were omitted from the census frame that was used to draw the sample of PSUs for the Master Sample. Other omissions from the frame included all institution EAs except workers' hostels, convents and monasteries. EAs in the census database that were found to have less than sixty dwelling units during listing were pooled.

The Master Sample is a multi-stage stratified sample. The overall sample size of PSUs was 3 000. The explicit strata were the 53 district councils. The 3 000 PSUs were allocated to these strata using the power allocation method. The PSUs were then sampled using probability proportional to size principles. The measure of size used was the number of households in a PSU as counted in the census.

The sampled PSUs were listed with the dwelling unit as the listing unit. From these listings systematic samples of dwelling units per PSU were drawn. These samples of dwelling units form clusters. The size of the clusters differs depending on the specific survey requirements. The GHS uses one of the clusters that contain ten dwelling units.

**UNITS OF OBSERVATION:** Households

**DATE OF DATA COLLECTION:** July 2007

**EXTENT OF DATA COLLECTION:** 4 data files in ASCII and machine-readable documentation.

**FILE SPECIFICATIONS:**

**Part 1: House Data**

Number of cases:	29 280
Number of records:	29 280
Number of records per case:	1
Logical Record Length:	406
Number of variables:	192
Number of kilobytes:	11 638 KB

**Part 2: Person Data**

Number of cases:	109 975
Number of records:	109 975
Number of records per case:	1
Logical Record Length:	141
Number of variables:	89
Number of kilobytes:	15 251 KB

**Part 3: Tourism Data**

Number of cases:	29 280
Number of records:	29 280
Number of records per case:	1
Logical Record Length:	326
Number of variables:	111
Number of kilobytes:	9 351 KB

**Part 4: Worker Data**

Number of cases:	74 856
Number of records:	74 856
Number of records per case:	1
Logical Record Length:	105
Number of variables:	45
Number of kilobytes:	7 530 KB

**PUBLICATIONS:**

Statistics South Africa. 2007. General Household Survey, 2007. Pretoria: Statistics South Africa.

**NOTES:**

Most questions in the General Household Survey questionnaire are pre-coded, i.e. there is a set number of choices from which one or more must be selected. For open-ended 'write-in' questions, the description will note that post-coding occurred and explain how this was done. Most variables have been pre-coded from the questionnaire and is not repeated in the variable description. Where the coding is not apparent, the description either provides the codes or indicates where code lists are to be found.

**RESTRICTIONS:**

Only available to bona fide researchers.

**DATE STUDY ADDED:**

January, 2008